

## CHOICE AND DELAY OF REINFORCEMENT<sup>1</sup>

SHIN-HO CHUNG AND R. J. HERRNSTEIN

HARVARD UNIVERSITY

Pigeons were trained to peck either of two response keys for food reinforcement on equated aperiodic schedules. The distribution of responding at the two keys was studied as reinforcement was delayed for various durations. The relative frequency of responding at each key was shown to match the relative immediacy of reinforcement, immediacy defined as the reciprocal of the delay of reinforcement.

The present experiment extends the investigation of reinforcement delay described by Chung (1965). In the earlier study, pigeons were trained to peck two response keys and received food reinforcement on equated variable-interval schedules. The reinforcements for responses on one key were delayed by various durations of time out, while the reinforcements for responses on the other were immediate. The findings suggested a negative exponential relation between the relative frequency of responding and the duration of the delay. The present experiment also used a concurrent procedure, but differed in setting delays of various durations for both of the response alternatives.

### METHOD

#### *Subjects*

Six male White Carneaux pigeons, experienced in a wide variety of experimental procedures, were maintained at approximately 80% of free-feeding weight.

#### *Apparatus*

A pair of experimental chambers for pigeons was used. Each chamber contained two response keys, spaced 9 cm apart, and a feeder providing 3-sec access to food for reinforced responses. Effective pecks had to be of at least 10-g force and each operated a relay to provide

auditory feedback to the pigeon. The chamber was illuminated by a white bulb, and, except when the magazine was operated, each response key was transilluminated by a 7-w red bulb. A continuous white masking noise was delivered during sessions.

#### *Procedure*

Pecks on either of the two response keys were, at first, reinforced on a variable-interval schedule with an average interval of 1 min. Two independent programmers arranged reinforcements for responses on the two response keys, with the restriction that a switch from one response key to the other prevented reinforcement for 1 sec (changeover delay or COD 1-sec). When the rate of pecking on the two keys became stable and approximately equal, delays of reinforcement were initiated. For subjects 237, 236, 415, and 416, responses on the left key were reinforced after an 8-sec delay (standard key). Reinforcements for responses on the right key were delayed for various durations ranging from 1 to 30 sec (experimental key). The intervals of delay imposed on the experimental key, in irregular order, were: 1, 2, 4, 6, 8, 12, 16, 20, 24, and 30 sec. For the other two subjects, 242 and 211, reinforcements for responses on the standard key were delayed for 16 sec. The intervals explored on the experimental key were: 2, 4, 6, 12, 20, 24, and 30 sec. From 21 to 40 sessions were given for each pair of delay intervals, depending on how quickly stable performance was attained. Sessions were terminated after the 60th reinforcement.

Between the response-to-be-reinforced and delivery of the reinforcement, the chamber

<sup>1</sup>This work was supported by grants from the National Science Foundation to Harvard University and from the National Institutes of Health (Grant M.H. 12108-01) to the Foundation for Research on the Nervous System. Reprints may be obtained from either author, Psychological Laboratories, William James Hall, Harvard University, Cambridge, Mass. 02138.

was darkened and responses produced no auditory feedback. As expected, this delay period contained virtually no responses. In Chung's previous study (1965), reinforcement-delay periods were produced by responses on one key only, while responses on the other produced an equal number of time-out periods uncorrelated with reinforcement. This feature of his procedure, which was an effort to cancel out the effects of time out *per se*, as distinguished from the effects of reinforcement delay, was not duplicated here. One subject (S-326), not listed above, was disqualified from the experiment after initial training, since it responded only on one of two keys when any pair of delay intervals was imposed. This may have been due to this subject's prior exposure to an experiment in which reinforcements were followed by blackouts for one of the two response keys and no reinforcements were followed by blackouts for the other.

### RESULTS

The relative frequency of responding on the experimental key was found to be a joint function of the delay intervals on that key and on the standard key. With an 8-sec delay on the standard key, the relative frequency of responding on the experimental key varied from 0.82 to 0.15 as the delay interval varied from 1 to 30 sec. With a 16-sec delay on the standard key, the relative frequency of responding on the experimental key varied from 0.91 to 0.34 as the delay interval varied from 2 to 30 sec. In Fig. 1, the relative frequency of responding on the experimental key is plotted against the duration of the delay. Points for each subject in Fig. 1, and in the subsequent figures, were obtained by averaging the performances of the final 10 sessions at each duration. The upper curves in Fig. 1 were obtained from the group receiving the 8-sec delay for the standard key; the lower curves are from the group with the 16-sec delay for the standard key.

The gross features of the two sets of curves are in agreement. In both instances, the general trend appears to be a monotonically decreasing function. However, there are certain conspicuous differences between the two sets of curves. First, at each value of delay, the height on the ordinate for the upper curves is, in almost every instance, less than the

height for the lower. In other words, a smaller fraction of responses occurs for a given delay if the alternative is 8 sec than if it is 16 sec. Secondly, the evidence for upward concavity is more pronounced for the upper curves than for the lower, where the evidence is questionable at best. And, finally, the upper curves appear to pass through the value of 0.5 on the

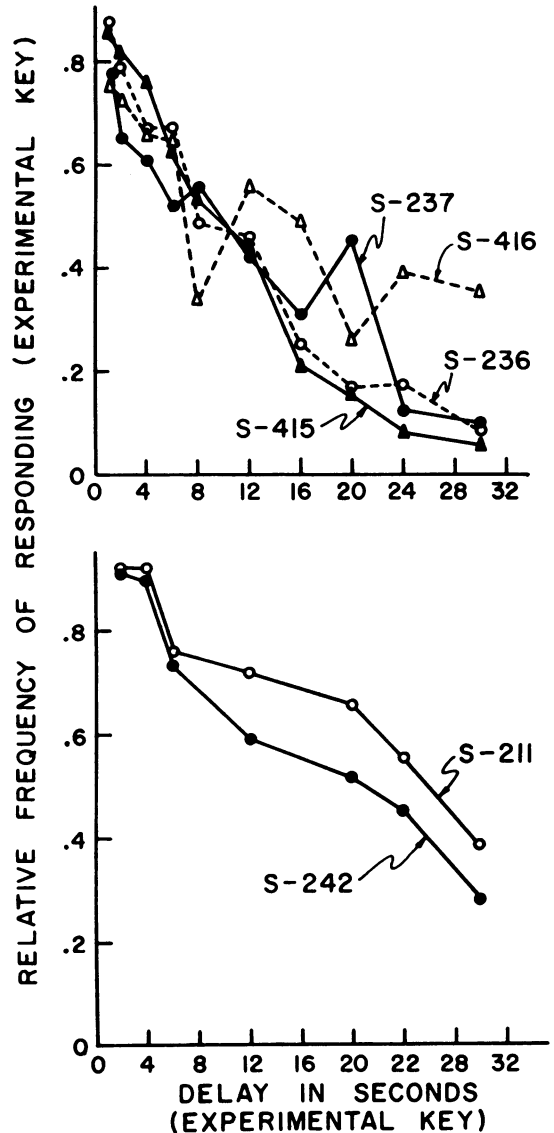


Fig. 1. Ratio of the number of responses on the experimental key over the total number of responses on both keys as a function of the duration of the reinforcement delay for the experimental key. Each point is the average of 10 sessions for one subject. The upper curves are for the subjects with an 8-sec standard delay; the lower curves for the subjects with a 16-sec standard delay.

ordinate at close to the expected value on the abscissa—8 sec—whereas the lower curves are clearly elevated about 0.5 at 16 sec. Assuming that the pigeon would distribute its responses equally between the two keys when delay durations for the two keys were equal, the elevation of the lower two curves suggests that some degree of artifactual key-preference was present for the 16-sec group.

These various features of the individual functions are shown for the averaged groups in Fig. 2, which also includes the data from Chung's 1965 study in which the pigeons (a group of three) were choosing between the delays shown on the abscissa and immediate reinforcement. The addition of this third function further substantiates the trends already noted. With presumably zero delay on the standard key, the degree of upward concavity is further accentuated. Moreover, this additional function is situated even lower on the ordinate than the other two. One further aspect of this added function might be noted. When responses on either key were reinforced immediately, *i.e.*, at zero on the abscissa, the pigeons responded equally, as would be expected. Contrary to expectation, however, the curve rises with a delay of 1 sec before it starts

to decline along the exponential curve used by Chung to describe these data. This rise is probably genuine, for it was observed for each of the three subjects. Nor is the rise actually contrary to intuition, considering the actual circumstance. A 1-sec delay of reinforcement gives the pigeon time to get its head into, or close to, the feeder opening so that it may start eating as soon as the feeder arm is within reach. The putative "immediate" reinforcement, in fact, involves whatever delay is accounted for by the pigeon lowering its head to the feeder and probably involves a shorter effective reinforcement duration than a reinforcement delayed for 1 sec. It may, therefore, be entirely proper to consider the second crossing of the 0.5 level, at about 1.5 sec of delay, as the point at which the delays for the two keys were actually equal.

The relative frequency of responding at the experimental key ( $\frac{R_e}{R_e + R_s}$ ) as a function of the relative duration of the delay intervals on the standard and experimental keys (*i.e.*,  $\frac{d_e}{d_e + d_s}$ , with  $d_e$  for the experimental delays and  $d_s$  for the standard delay) is shown in Fig. 3. The data obtained from the earlier study (Chung, 1965) were included in Fig. 3 by taking a small constant as the actual delay inter-

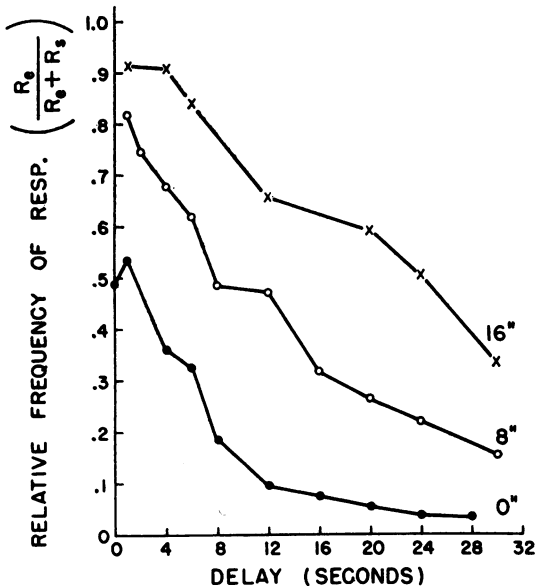


Fig. 2. Relative frequency of responding on the experimental key as a function of the reinforcement delay for that key, averaged across subjects. The parameters refer to the duration of the standard delay. The bottom curve was taken from Chung (1965).

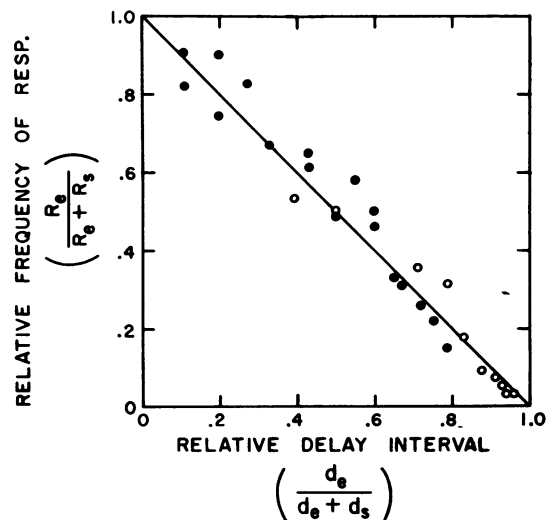


Fig. 3. Relative frequency of responding on the experimental key as a function of the ratio of the duration of the delay for that key over the sum of the durations of the delays for the two keys. The filled circles are values for the 8-sec and 16-sec groups separately; the open circles are from Chung (1965).

val for what is nominally immediate reinforcement. The value of the constant (1.6 sec) was estimated so as to minimize the sum of the squared deviations between the observed values and the function predicted from the present findings. These earlier data are shown as open circles, presenting the average of three pigeons. The filled circles show separately the 8-sec and the 16-sec groups in the present experiment. Figure 3 shows that the relative frequency of responding closely matches the relative immediacy of the delay intervals, if relative immediacy is taken as the complement of relative delay. The more familiar increasing diagonal could just as well have been obtained using immediacy, defined as the reciprocal of delay, as the independent variable. Thus,

$$\frac{R_e}{R_e + R_s} = \frac{d_s}{d_s + d_e} \quad \text{or} \quad (1)$$

$$\frac{R_e}{R_e + R_s} = \frac{i_e}{i_e + i_s}, \quad (2)$$

where  $i$  is the reciprocal of  $d$ , and the subscripts distinguish the standard key from the experimental key. Although the data approximate the diagonal reasonably well, the 16-sec group tended to fall consistently above the diagonal, as would readily be predicted from the average curve in Fig. 2. This deviation seems to be attributable to a key preference

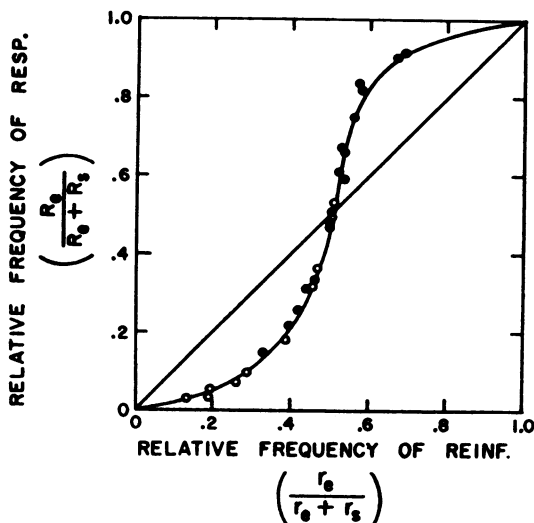


Fig. 4. Relative frequency of responding on the experimental key as a function of the relative frequency of reinforcement associated with that key. Filled circles are for the 8-sec and 16-sec groups separately, open circles are from Chung (1965).

since only the intercept, and not the slope, of the function is affected.

Although the programmed rate of reinforcement for responses on the two keys was identical, the relative frequency of reinforcement actually delivered for responses on each varied systematically as a function of the relative delay interval because rate of responding decreased with increases in the delay interval. Figure 4 shows the relative frequency of responding as a function of relative frequency of reinforcement actually delivered for responses on that key. Once again, the open circles show the data from Chung (1965). The function deviates systematically from the linear relation, indicating that the changes in the relative frequency of reinforcement do not sufficiently account for changes in the relative frequency of responding, and that delay itself is instrumental for the function in Fig. 3.

The absolute rates of responding, as well as the reciprocal relation between the responding on the two keys, are shown in the three sections of Fig. 5, corresponding to the three values of the standard delay studied in Chung (1965) and the present experiment taken together. The filled circles show the average rate of pecking at the key with the varying duration of delay; the open circles, at the key with the fixed duration as indicated for each section. The curves were fitted by eye. Various features of these curves are readily noted. It is clear that a change in the delay value for one key affects the rate of responding on both keys; that the effect is opposite and, to some fair degree of approximation, symmetrical; that the amount of curvature in these functions decreases with an increasing standard delay, and that the point of intersection for each pair of curves is further to the right as the standard delay is increased. All of these features, as will be shown, are characteristic of behavior that obeys the matching relation depicted in Fig. 3.

## DISCUSSION

The central fact disclosed by the present study is that the relative frequency of responding matches the relative immediacy of reinforcement in a two-response situation. Hence, delay of reinforcement may now be added to those other variables, like frequency and amount of reinforcement, whose effects are

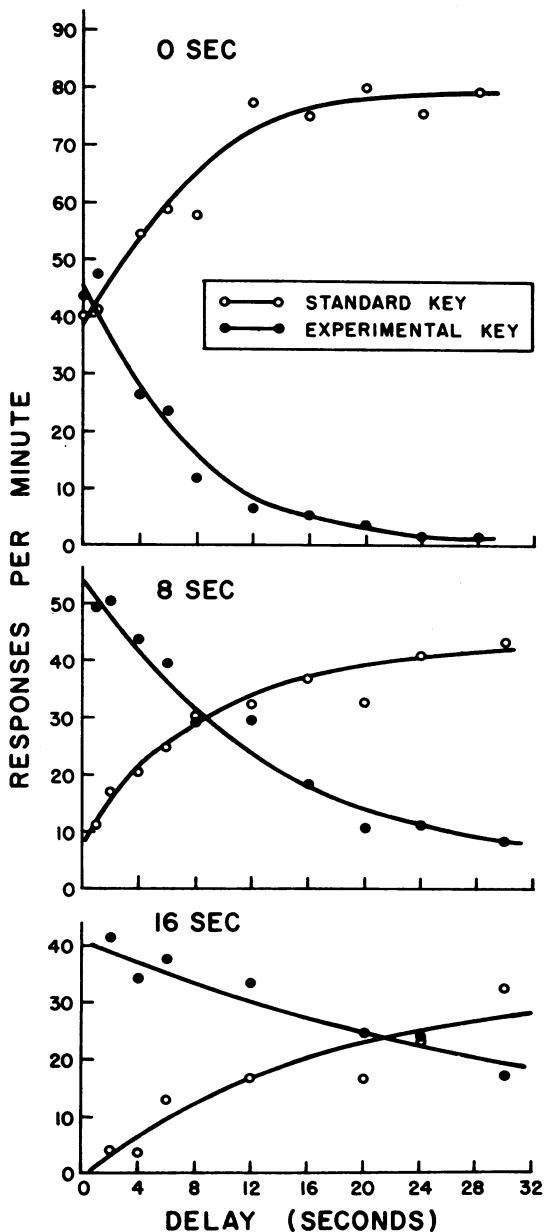


Fig. 5. The absolute rate of responding on the two keys as a function of the delay for the experimental key, averaged across subjects. The three sections are for the three durations of the standard delay. Filled circles show rates for the experimental key; open circles, for the standard key.

adequately depicted as the subject's matching of its behavior to the ratio of the magnitudes of the independent variable. Thus, the ratio of choices in a two-choice situation has been shown to equal the ratio of rates of reinforcement (Herrnstein, 1961); to equal the ratio of

amounts of reinforcement (Catania, 1963); and, here, to equal the ratio of immediacies of reinforcement. But since Catania's experiment varied amount of reinforcement by varying the duration of feeder operation in a conventional key-pecking apparatus, it could be argued that these three findings arise from the properties of the pigeon's perception of time and nothing more. Baum (1966), however, demonstrated essentially the same relativity of choice using rats instead of pigeons and sucrose concentration—in no obvious way a temporal parameter—as the independent variable.

The traditional interest in delay has not, however, been in its effects on choice, but on acquisition and its asymptote. Although Hull (1943) used the data from choice situations, such as Perin's (1943) and Anderson's (1932), he treated them as if they revealed only the habit strength of the response whose reinforcement was being delayed and not the operation of a choice mechanism. From these data, Hull concluded that the asymptotic habit strength of a response was a negative exponential function of the delay of reinforcement. At least superficially, Chung's recent study (1965) confirms Hull's hypothesis. As stated earlier, Chung studied, in a two-choice procedure for pigeons, the effects of delaying reinforcement for one of the two alternative responses, while the other alternative was always reinforced immediately. As the duration was varied it was found that the effects could be adequately summarized by a decreasing exponential function between the relative frequency of responding on the key for which reinforcement was delayed and the duration of the delay. This agreement with Hull is all the more impressive for having been based on results from different species—pigeons *versus* rats—and from a number of different types of apparatus, from the simple T-maze to the two-key pigeon box. The agreement notwithstanding, certain problems remain unsolved.

There is, first, the fact that in picking the exponential function Hull was relying heavily on Anderson's results, obtained from an old-style discrimination-box procedure in which rats chose one of either of two or four compartments presented simultaneously. Any choice was rewarded with food, but, depending upon the compartment chosen, the rat was detained for periods of 1 to 4 min. A rereading of Anderson's original paper shows that although it

is clear that at some point in training the rats were distributing their choices in approximate agreement with an exponential function, his data suggest that with further training the rats were tending to choose only the compartment with the shortest delay. Such a tendency is hardly surprising, for there is nothing in the situation to dispose a rat to choose a longer delay when a shorter one is available and equally profitable. That situation existed in Anderson's study, but not in the present experiment, as is shown below. In his curve-fitting, Hull used only the data from the earlier point in training, and spoke of the distribution of choices as if it were dictated by the rats' inability to form discriminations between certain short delays and longer ones, thereby tacitly making the effect of delay one of discriminability rather than of reinforcing power.

Whatever the answer may be as regards Anderson's study and the problem of temporal discrimination in the rat, the parallel to Chung's study is obviously questionable. Since the later study used a pair of variable-interval schedules, with delays of reinforcement for one of the choices, there was a factor disposing the pigeons to respond to both alternatives. Unlike the continuous reinforcement in Anderson's experiment, and, incidentally, in virtually all of the other experiments in the preoperant literature, the aperiodic schedule may impose a penalty in reinforcements lost should the animal respond exclusively to either alternative, the one with the shorter delay or otherwise. Chung's exponential, then, may well have been describing an asymptotic performance, but its agreement with Anderson's pre-asymptotic data may be devoid of substantive meaning.

This is not the only problem. Hull noted that a number of experimenters had examined the effects of reinforcement delay and had, contrary to their own expectations, found virtually none for delays that, in other studies, proved to be more than ample. John B. Watson was one of these, having varied delay of reinforcement in his version of the puzzle box (1917) and found that his rats were unaffected by delays up to 30 sec. More recently, Ferster (1953) argued that delays of even longer durations can be bridged with no decrement in performance if the animal is equipped with the suitable behavior to medi-

ate the temporal gap. Ferster was, in effect, agreeing with Hull, who also attributed the lack of potency of delays to the interfering effects of secondary reinforcement. And more recently still, Logan (1960) examined reinforcement delay in a runway for rats and found relatively minor quantitative effects of delays between 1 and 30 sec.

Finally, the present study, in which the pigeons were choosing between pairs of delays, suggests, as will be shown, that the exponential function may not be the best summary of the effects of delay, the earlier agreement notwithstanding. The data in Fig. 2 are replotted in Fig. 6 to test the validity of the negative exponential function as a description of the results. The ordinate is the relative frequency of responding on a logarithmic scale; the abscissa is the duration of the delay on a linear scale. Negative exponentials plot as straight lines with negative slope in such a semi-logarithmic coordinate. The three straight lines shown were drawn by eye through the data points of the present experiment and of Chung's earlier study. The parameters on the curves give the duration of the standard delay for each of the three conditions.

In the absence of other considerations, the fit between data and theory in Fig. 6 would undoubtedly enhance the credibility of both. Except for a slight tendency toward upward concavity, most evident for the bottom function, the points hover close to the straight lines called for by the hypothesis of a negative exponential. The data used by earlier workers to substantiate the negative exponential rarely fit so well, over so broad a range, with so little averaging necessary. The other consideration,

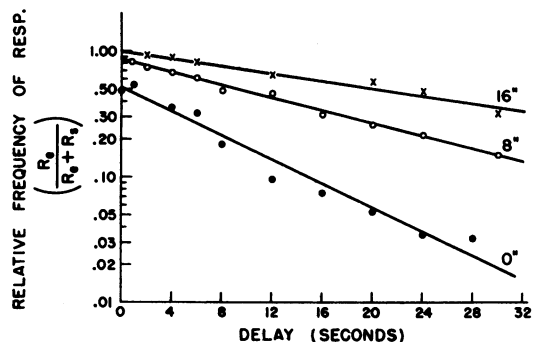


Fig. 6. Relative frequency of responding on the experimental key, on a logarithmic scale, as a function of the delay for that key, averaged across subjects. The parameters refer to the duration of the standard delay.

however, that makes the negative exponential suspect, is the function in Fig. 3, which shows that the relative frequency of responding matches the complement of the relative duration of the delay. It is readily shown that this matching relation is mathematically incompatible with the negative exponential suggested by Fig. 6. No proof will be given since it is virtually self-evident that the two following equations, representing the matching rule and the exponential rule, respectively, cannot both hold:

$$\frac{R_e}{R_e + R_s} = \frac{d_s}{d_e + d_s} \quad (3)$$

$$\frac{R_e}{R_e + R_s} = ka^{-bd_s} \quad (4)$$

in which  $k$ ,  $a$ , and  $b$  are constants and the other symbols are as used above.

Two incompatible theories can fit a single set of data points only if there is enough variability in the data to provide the requisite level of ambiguity. The present data clearly provide at least this level. Fortunately, however, the quantitative nature of the two theories permits one further step in the analysis. If Equation 3 did, in fact, correctly describe the data, the points would deviate from the straight lines in Fig. 6 in tending to be concave upward. Moreover, the upward concavity would be greatest for the bottom curve and so slight as to be virtually undetectable for the upper curve, with the middle one falling between. This mathematical implication of the matching function seems to be borne out by the pattern of deviations of the data from the straight lines. On the other hand, the contrary assumption, that Equation 4 is correct, predicts that the points in Fig. 3 would tend to deviate from the diagonal by being concave downward, with the discrepancy between predicted and obtained values increasing at both ends of the function. This suggestion appears to account for none of the variance of the points around the theoretical line.

The present experiment was not designed as a test of the two formulations being discussed. It is not surprising, then, that a firm conclusion is not forthcoming. The data slightly favor the matching relation, but not enough to exclude the exponential. On the other hand, it is now clear that the two theories, in spite of their mutual incompatibility,

predict similar relative frequencies of responding within the range of delays usually studied. To separate between the theories quantitatively would require an experiment in which very long and very short delays would be directly compared in a choice procedure; it is at these extremes that the two theories diverge measurably.

There are, however, non-quantitative arguments that favor the matching rule. The theoretical curve in Fig. 3, which asserts that a given ratio of delays will produce a given ratio of responses, independent of the absolute levels involved, is the sole theoretical line allowed by the matching hypothesis. In contrast, the exponential hypothesis allows any negative slope and intercept in the semi-logarithmic coordinate used in Fig. 6. The matching hypothesis, in other words, is by far the more restrictive and powerful of the two. This is also evident from the fact that Equation 3 calls for no free parameters, while Equation 4 calls effectively for two.

The approximate symmetry of the rising and falling curves in Fig. 5 means that the number of pecks on the two keys summed together maintained an approximate constancy, in spite of the changing duration of delay for one of the keys. A comparable insensitivity of responding to the duration of delay was already noted in connection with experiments by Watson (1917), Ferster (1953), and Logan (1960). The results of these experiments, and the summed rate of responding in the present experiment, have one thing in common: all are based on absolute measures of performance. The results obtained by Anderson (1932), Perin (1943), and Chung (1965), on the other hand, are based on relative measures. In the few studies that show large effects of delay and that use absolute measures, *e.g.*, Skinner (1938, p. 139 ff) or Dews (1960), the animal's response postpones reinforcement. This is a direct contingency favoring a low rate of responding and not reinforcement delay in the usual sense of the term. Delay of reinforcement, then, like frequency of reinforcement (Herrnstein, 1961) or amount of reinforcement (Catania, 1963), is a variable that influences more the relative strength of a response among a set of response alternatives than the absolute strength of a response in isolation. Given this insensitivity of the over-all rate of responding in the present experiment, the various features

of Fig. 5 follow from the matching relation depicted in Fig. 3. It then follows, for example, that a change in delay duration on one key will alter in opposite directions the rates of responding on both keys, and that the functions become progressively less curved with increasing standard delays. For all its simplicity, the matching hypothesis appears able to account for the present results, and also for those to be found in the experimental literature.

### REFERENCES

- Anderson, A. C. Time discrimination in the white rat. *J. comp. Psychol.*, 1932, 13, 27-55.
- Baum, W. M. Choice in the rat as a function of probability and amount of reward. Unpublished doctoral dissertation, Harvard University, 1966.
- Catania, A. C. Concurrent performances: A baseline for the study of reinforcement magnitude. *J. exp. Anal. Behav.*, 1963, 6, 299-300.
- Chung, S. H. Effects of delayed reinforcement in a concurrent situation. *J. exp. Anal. Behav.*, 1965, 8, 439-444.
- Dews, P. B. Free-operant behavior under conditions of delayed reinforcement. I. CRF-type schedules. *J. exp. Anal. Behav.*, 1960, 3, 221-234.
- Ferster, C. B. Sustained behavior under delayed reinforcement. *J. exp. Psychol.*, 1953, 45, 218-224.
- Herrnstein, R. J. Relative and absolute strength of response as a function of frequency of reinforcement. *J. exp. Anal. Behav.*, 1961, 4, 267-272.
- Hull, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
- Logan, F. A. *Incentive*. New Haven: Yale University Press, 1960.
- Perin, C. T. The effect of delayed reinforcement upon the differentiation of bar responses in white rats. *J. exp. Psychol.*, 1943, 32, 95-109.
- Skinner, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
- Watson, J. B. The effect of delayed feeding upon learning. *Psychobiology*, 1917, 1, 51-59.

*Received August 30, 1966*